Brazilian version of the OSTRC Questionnaire on health problems (OSTRC-BR): translation, cross-cultural adaptation and measurement properties

Rafael Mirante Pimenta, Luiz Hespanhol, Alexandre Dias Lopes

Abstract

Background: The Oslo Sports Trauma Research Center - Questionnaire on Health Problems (OSTRC-QHP) has been shown to be a valid and reliable questionnaire for monitoring athletes with acute, overuse, and illness problems.

Objectives: To translate and cross-culturally adapt the OSTRC-QHP into Brazilian-Portuguese (OSTRC-BR), and investigate the measurement properties of this version.

Methods: The original (Norwegian) questionnaire was translated and culturally adapted following existing guidelines. Measurement properties of the OSTRC-BR were investigated in 134 athletes followed for 13 weeks. Reliability was analyzed via internal consistency and test-retest reliability. Responsiveness was assessed by effect size (ES) estimations, correlation of the OSTRC-BR severity score with the Global Perceived Effect scale (GPE) and the Numeric Pain Rating Scale (NPRS), and by the Receiver Operating Characteristic (ROC) Curve test based on the area under the curve (AUC).

Results: The OSTRC-QHP has been translated and cross-culturally adapted successfully. The Cronbach's alpha was 0.93 (95% CI: 0.92, 0.94). Intrarater reliability (ICC2,1) of 0.96 (95% CI: 0.94, 0.98) was classified as excellent. ES estimation showed large effects (1.14 to 1.39). OSTRC-BR correlation was moderate (0.58 to 0.72) with the GPE and substantial (0.84 to 0.89) with the NPRS. The AUC values were above 0.70 when analysing all types of health problems.

Conclusions: The OSTRC-BR presented adequate measurement properties. The OSTRC-BR is a valid, reliable, responsive, and, therefore, an adequate tool for monitoring Brazilian athletes from different sports during training and competitions.

© 2021 Associação Brasileira de Pesquisa e Pós-Graduação em Fisioterapia. Published by Elsevier España, S.L.U. All rights reserved.

KEYWORDS
Clinimetric; Rehabilitation; Responsiveness; Sport injuries surveillance
Introduction

The Oslo Sports Trauma Research Center—Questionnaire on Health Problems (OSTRC-QHP) was developed by Clarsen et al.\(^1\) to be a general sports injury surveillance system useful for measuring the prevalence, incidence, and severity of sports injuries over time. In comparison with standard methods used in several Olympic Games and World Championships, where injuries were registered based on medical attention or time-loss definition, the OSTRC-QHP allows the use of a broader injury definition like “any complaint”, providing a better understanding and information on the consequences of injuries.\(^1-4\) The OSTRC-QHP advantage is the continuous monitoring of health problems regardless of training or competition contexts and regardless of any severity threshold.

The OSTRC-QHP has been translated and adapted into Dutch, Danish, German, and Japanese languages.\(^5-8\) Although the Brazilian-Portuguese (BrPt) speaking populations have been the target of many health studies, no valid and reliable sport injury-monitoring tool is available in BrPt. Therefore, the objectives of this study were (1) to translate the OSTRC-QHP into the BrPt language; (2) to culturally adapt the translated OSTRC-QHP to the Brazilian context; and (3) to evaluate the measurement properties (content validity, reliability, responsiveness, and floor and ceiling effects) of the BrPt OSTRC-QHP version (i.e., OSTRC-BR).

Methods

Athletes and recruitment

The Olympic Training and Research Center (COTP) is a training center for athletes with about 1000 registered athletes. A convenience sample of 152 athletes based on the criteria by Terwee et al.\(^*\) for performing measurement property analysis (validation, reliability, and responsiveness) were recruited from the COTP in São Paulo, Brazil. The eligibility criteria were: (1) age between 13 and 30 years; (2) participant of the COTP training program; and (3) able to read, write, and understand BrPt. The following sports were selected as sources for athletes: basketball, handball, volleyball, football, swimming, and judo. Some sports were excluded due to the low ages of the athletes (gymnastics) and due to incompatible training schedules for data collection (athletics). For all athletes under the age of 18, written parental consent was required for participation in this research project.

Description of the measures

Three measurement tools were used in this study. The first was the OSTRC-QHP, which consists of four key questions recording any health problem that has altered an athlete’s sports participation, training volume, or sports performance, as well as the level of the related symptoms.\(^1\) The responses to each of the four questions are allocated a numerical value from 0 to 25. Questions 1 and 4 are scored 0–8–17–25, and questions 2 and 3 are scored 0–6–13–19–25. These are added to calculate a severity score (SS) from 0 to 100 for each health complaint. In the presence of an injury or disease, athletes needed to answer eight additional questions to inform the anatomical location of injuries and list illness symptoms as well as to provide additional information about the problem such as the number of days of time loss, and whether they have received medical attention. In case of multiple health problems, the questionnaire returned to the four key questions and repeated itself for each subsequent problem reported. The question number 5 of the OSTRC-QHP, classified the health problem as injury or illness. Injuries were any health problem related to musculoskeletal disorders or concussions, and illnesses were problems involving other body systems, such as respiratory, digestive, and neurological diseases, as well as nonspecific generalized psychological/social-related problems.\(^1\) Injuries were subcategorized as acute or overuse injuries. Acute injuries were defined as those whose specific onset event was clearly identifiable.\(^1\) Overuse injuries were those when a clearly defined onset event could not be identified.\(^1\)

The second measurement tool was the Numeric Pain Rating Scale (NPRS), which ranges from 0 (no pain) to 10 (worst possible pain).\(^10\) The NPRS was used as an external criterion for measuring pain intensity based on the last seven days. The third was the Global Perceived Effect (GPE) scale, an 11-point scale that ranges from –5 (vastly worse) through 0 (no change) to +5 (complete recovery).\(^11\) The question used on GPE was: “Compared to the beginning of this health complaint, how would you describe your health problem today?”. The GPE was used as an external criterion for measuring overall change. The GPE and NPRS have been cross-culturally adapted into BrPt and their measurement properties tested in previous studies.\(^10-12\)

Study design

The study was divided into two phases. Phase 1 consisted of translation, cross-cultural adaptation, and content validity assessment by 30 athletes. Phase 2 included responsiveness, floor and ceiling effects, reliability evaluation through a 13-week longitudinal study design. The athletes completed the OSTRC-BR, NPRS, and GPE once a week during the training session.

This was a longitudinal and prospective cohort study approved by the Research Ethics Committee of the Universidade Cidade de São Paulo (UNICID), São Paulo, Brazil (#058721/2016). This study followed the COnsensus-based Standards for the selection of health Measurement InStruments (COSMIN).\(^13\) The study phases and questionnaire logic are summarized in Fig. 1.

Translation, cross-cultural adaptation, and content validity

First, permission to translate and cross-culturally adapt the questionnaire was obtained from the first author of the OSTRC-QHP original version.\(^1\) Thus, the original Norwegian version was translated into BrPt according to Beaton et al.\(^14\) This process consisted of the translation of the items by two independent bilingual translators, both native BrPt speakers (one a physical therapist and the other without any medical problem).
training) who were also fluent in the Norwegian language. The first BrPt version was independently back-translated to Norwegian by two native Norwegian speakers who had no knowledge of the original tool and were not aware of the objectives of the study. The Norwegian and the BrPt versions were revised by the expert committee consisting of methodologists and health professionals with experience in conducting translation and cross-cultural adaptation studies, two bilingual rehabilitation specialists, and the forward and backward translators. The expert committee considered the original version and the final version similar and also were allowed to make changes to ensure semantic, idiomatic, and conceptual equivalence. The pre-final version of the OSTRC-BR was tested for comprehensibility and content validity with a sample of 30 athletes from the COTP. For the purpose of evaluating comprehensibility, athletes were asked at the end of this process about the difficulties or incomprehensibility of each question. In addition, an open field was provided at the end of each questionnaire for any comments. To establish the content validity of the questionnaire, athletes and the expert committee (clinicians and sports injury epidemiologists) were involved in this step. Athletes’ interpretations coupled with rational analysis by the clinicians and sport injury epidemiologists related to the project were considered. Athletes’ interpretations were assessed by adding a “not applicable” alternative to all BrPt questions. In addition, we informed all athletes that if they judged that the content of the questionnaire was not applicable to their cases, they should choose the “not applicable” option and provide the reason in the open field available.

Measurement properties

Reliability was analyzed via internal consistency and test-retest reliability. Internal consistency (IC) is the extent of interrelatedness among the items of a scale. The IC was calculated using Cronbach’s alpha, with a value between 0.70 and 0.90 considered to be adequate. We hypothesized a Cronbach’s alpha estimate higher than 0.90 based on previous studies. The test-retest was conducted at week 10, with at least 50 randomly selected athletes who reported health complaints (injuries/illnesses). The retest was performed 48/72 h after the first response. The Intraclass Correlation Coefficient (ICC2,1) values were classified as low (< 0.40), moderate (0.40–0.75), substantial (0.76–0.90), or excellent (> 0.90). We hypothesized an ICC value >0.75 based on previous studies.

To assess responsiveness, the OSTRC-BR Severity Score (SS) was compared to the global perceived effect (measured by the GPE scale), and pain intensity (measured by the NPRS scale). Responsiveness was assessed for all cases and for each type of health problem (acute/overuse/illness). Acute and overuse injuries responsiveness was measured by correlating the change score (score at the sixth week minus baseline score) from the SS with the change score of the NPRS.
and the score of the GPE (at the sixth week from the first injury complaint record). Illness responsiveness was measured by correlating the change score (score at the second week minus baseline score) from the SS with the change score of the NPRS and the score of the GPE (at the second week from the first illness complaint record). We used different time intervals for injuries and illnesses based on the average duration presented from previous studies. A correlation < 0.40 was classified as low, 0.40–0.75 as moderate, 0.76–0.90 as substantial, and > 0.90 as excellent. We hypothesized a low correlation between the OSTRC-BR and GPE, and a low correlation between OSTRC-BR and NPRS, since GPE and NPRS measure changes in different (but health related) constructs when compared to OSTRC-BR.

Responsiveness was also assessed by calculation of effect sizes (ES). ES was estimated using the mean OSTRC-BR change score (week when the problem was registered and the sixth week for acute and overuse injuries or week when the problem was registered and second week for illnesses) divided by the standard deviation at the first week the health complaint was recorded. ES was calculated with a confidence interval of 84% of the sample, because non-overlapping 84% confidence intervals are equivalent to a Z scores at a 0.05 level. A score of less than 0.20 indicates a slight change, 0.50 indicates a moderate change, and 0.80 or greater indicates a large change. Our hypothesis was that the ES of the OSTRC-BR would be higher for acute injuries than for overuse injuries due to different injury patterns and a higher severity score of acute injuries reported in previous studies.

The Receiver-Operating-Characteristic (ROC) Curves were calculated to show the probability of correct discrimination between patients who improved from those who did not improve and between patients who worsened from those who did not worsen. The OSTRC-BR change scores were classified as a dependent variable for all analyses. The GPE scores were dichotomized to classify the participants as “improved” or “not improved” based on a cut-off score of 3 points or higher being considered as improved, and patients with scores between 0 and 2 were classified as “not improved”. For the analyses of patients who worsened, the GPE scores were dichotomized into “worsened” or “not worsened” based on a cut-off score of –3 points or lower being considered as worsened and patients with scores between –2 and 0 being classified as “not worsened.” The same analysis was done with the NPRS, an improvement equal or greater than 30% between initial assessment and follow-up assessments was categorized as “improved” and growth between 0 and 29% being classified as “not improved”. To analyze patients who worsened a change equal or greater than 30% between initial assessment and follow-up assessments was categorized as “worsened” and a change between 0 and 29% was classified as “not worsened”. This analysis was based on the area under the curve (AUC), and values of 0.70 or higher were considered to be responsive. Our hypothesis was that the AUC estimate for the OSTRC-BR would be higher for acute injuries than for overuse injuries, based on injury patterns and the challenges related to measuring overuse injuries in sports.

Floor or ceiling effects were calculated and considered to be present if more than 15% of the sample reached these values. We expected the presence of a floor effect because we hypothesized that many athletes would not present health problems during the follow-up period.

The normality of the data was inspected using histograms, and descriptive analyses were presented for the sample. All the analyses were performed using SPSS Version 22 (SPSS Inc, Chicago, IL).

Results

Translation and cross-cultural adaptation

The participants’ demographics are summarized in Table 1. The following adaptations were made by the expert committee; changed the term “health problems” with the terms “difficulties/discomfort” in all answers of the first question, and replace the term “injury” with the term “discomfort” in the question statement 2, 3, and 4 of the questionnaire. Another change was made in the last answer option of questions 2 and 3, where the answer “cannot participate at all” was replaced with “I cannot train or compete at all due to discomfort/illness or health problems”. The last change was made for question 11 (contact with the medical team), where an alternative response referring to the option of reporting the complaint to “other people” was included.

During the pre-test 30 athletes answered the four main questions, 19 of whom presented with health problems and answered the eight descriptive questions of injuries or illness. No difficulty or comprehensibility issues in answering the questionnaire were reported, as no athlete chose the “not applicable” option. Thus, the content validity of the OSTRC-BR was judged to be satisfactory, and the final version of the OSTRC-BR was established (Fig. 2).

Measurement properties

A total of 134 athletes were followed for 13 consecutive weeks, and 1829 questionnaires were completed during this period. The participants’ demographics are summarized in Table 1 and the results of the measurement properties analyses are summarized in Table 2. The response rate of the weekly OSTRC-BR was 94% (95% CI: 93, 95). In total, 118 athletes reported 735 health problems, of which 656 (89%) were classified as injuries (388 overuse and 268 acute injuries) and 79 (11%) as illnesses. The reliability analysis (Table 2) yielded a Cronbach’s alpha value of 0.93 (95% CI: 0.92, 0.94). The test-retest (n = 50) was classified as excellent with an ICC2,1 of 0.96 (95% CI: 0.94, 0.98).

The responsiveness analysis was performed based on 313 health problems that were reported by 92 athletes, consisting of 128 acute injuries, 106 overuse injuries, and 79 illnesses. As summarized in Table 2, responsiveness analysis showed a moderate (0.58–0.72) correlation between the change scores on the OSTRC-BR with the GPE scores for all types of health problems and a substantial (0.84–0.89) correlation with the NPRS score (Table 2). The ES to analyze responsiveness showed large effects (above 0.8) considering a 6-week period for injuries and a 2-week period for illnesses. The ES for OSTRC-BR was 1.14.
The OSTRC-BR AUC values were above 0.70 when analysing all types of health problems.

The OSTRC-BR presented a considerable floor effect because 60% (n = 1094) of all questionnaires (n = 1829) were completed without reporting a health problem.

Discussion

This was the fifth study that translated the OSTRC-QHP, evidence of the importance of this instrument for clinical and scientific practice in sport sciences. However, only this study evaluated the responsiveness of the OSTRC-QHP. We performed the responsiveness analysis to investigate the ability of the OSTRC-BR in detecting changes over the time. We believe that the fluctuation of the SS may reflect changes in athlete’s health status and has to be responsive to changes in other instruments that measure health status, such as the GPE scale or NPRS scale.

Translation and cross-cultural adaptation

The process used in this study was similar to the process used in the Danish and German translations of the OSTRC-QHP. A semantic adaptation was made to the final version of the OSTRC-BR because the expert committee understood that some athletes might not classify some health complaints as “injuries” due to the BrPt semantics of the term. In BrPt, the literal translation of the term “injury” has a negative connotation, which can reduce athletes’ likelihood of reporting minor health problems. This negative connotation was also noted in the Japanese translation. It is also important that the adaptation of an instrument possesses a cultural fit, that is, its preparation for use in different cultural contexts. Examples of these adaptations are that the German authors added a question on weekly training volume to obtain the exposure data of the participants, whereas in the Danish version, an additional “no complaint” alternative was included in the questionnaire to simplify things when no injuries or illnesses were reported in questions 1–4.

Clinimetric properties

To assess the measurement properties of the OSTRC-BR, we followed the COSMIN recommendations. The Cronbach’s alpha presented a mean value of 0.93. According to Terwee et al., a Cronbach’s alpha value higher than 0.90 is indicative of the redundancy of items. However, the Cronbach’s alpha presented a mean value of 0.93.
alpha didn’t improve with the item deleted option. Nevertheless, these findings are consistent with the results of the original Norwegian (0.96), German (0.92), Danish (0.90) and Japanese (0.96) versions.\textsuperscript{1,6–8}

The test-retest reliability was excellent with a mean ICC\textsuperscript{2,1} value of 0.96. In the original (Norwegian) version, the reliability was not evaluated. In the German version, the ICC\textsuperscript{2,1} was 0.91, with the same test-retest interval (i.e., 48/72 h).\textsuperscript{1,8} In the Danish version, the ICC\textsuperscript{2,1} was 0.72 when the test-retest interval time was one week, and it was 0.62 when the interval was two weeks.\textsuperscript{7} A longer interval between the test-retest, as used in the Danish version, may reflect natural fluctuations in the severity of injuries and illnesses, which may result in a lower ICC.\textsuperscript{1,8} The higher ICC (0.97) found in the Japanese version may be explained by the 24/72 h interval.\textsuperscript{6}

The challenge of establishing responsiveness is related to the need of having a gold standard instrument of change to compare with the instrument under investigation. According to de Vet et al.\textsuperscript{15} when there is no gold standard comparator (construct validity approach) it is necessary to formulate and test specific hypotheses a priori to evaluate
Our results showed that all responsiveness hypotheses were confirmed.

The correlation tests between the OSTRC-BR change scores with the GPE scores and NPRS change scores were positively classified as moderate and substantial, respectively. This was expected as according to de Vet et al. \textsuperscript{15} in responsiveness studies low to moderate correlations are often found when analyzing agreement between the change scores of different instruments. A high correlation would not be expected because the constructs that these instruments evaluated are not the same even though all are related to health status. \textsuperscript{15} Other ways to assess responsiveness in the construct approach are the ES and AUC analysis.
Table 3  Average weekly and cumulative severity score of health problems.

<table>
<thead>
<tr>
<th></th>
<th>Overuse injury (n = 388)</th>
<th>Acute injury (n = 268)</th>
<th>Illness (n = 79)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average weekly prevalence</td>
<td>26 (24, 28)</td>
<td>12.1 (11.0, 12.0)</td>
<td>4.2 (3.3, 4.5)</td>
</tr>
<tr>
<td>Average weekly severity score</td>
<td>30 (30, 31)</td>
<td>44 (42, 46)</td>
<td>43 (39, 47)</td>
</tr>
<tr>
<td>Cumulative severity score</td>
<td>1042 (965, 1118)</td>
<td>693 (561, 824)</td>
<td>247 (157, 336)</td>
</tr>
<tr>
<td>Total lost training days</td>
<td>138</td>
<td>283</td>
<td>111</td>
</tr>
</tbody>
</table>

Prevalence are presented as percentage (95% confidence interval). Severity score data are presented as total value and 95% confidence interval.

Based on previous studies, who showed that the SS of acute injuries were higher than overuse injuries, we expected that ES and AUC values of acute injuries would be higher than overuse injuries.1,4 As hypothesized, the ES was higher for acute injuries. Measures to assess overuse injuries are really a challenge for sports medicine. Possibly the pattern of lower SS on the OSTRC-BR and lower NPRS score, when compared to acute injuries and illnesses, reflects a lower change score, which in turn justifies a smaller ES for all measures of overuse. The responsiveness assessed by the AUC was higher than 0.70 for all types of health problems when compared with the GPE and NPRS dichotomized scores for patients who improved/not improved and worsened/not worsened. All this responsiveness results provide important information that the OSTRC-BR can measure changes through the analysis of the increase or decrease of the SS.

We suggest that future studies should focus on monitoring a specific population (i.e., athletes) with a specific type of injury (i.e., knee tendinopathy), correlate the OSTRC SS with a gold standard tool, and compare the findings with our results.

Limitation

This is the first study to investigate responsiveness of the OSTRC-QHP, so this analysis was a critical part of the current study. Therefore, some limitations should be considered. First, the correlation analysis was made by subtracting the SS. No previous study has done this analysis and such calculation can influence the result of the severity of the health problem, because OSTRC-QHP questions have different scores for each question. However, because responsiveness is a paramount measure property to be considered by researchers aimed at investigating the effectiveness of interventions in decreasing the risk and/or the number of sports injuries, as the sports injury prevention randomized controlled trials, we hope that our study will serve as a stimulus for further investigation and enhancements in the methods on the responsiveness analysis of the OSTRC-QHP in its different versions. Second, we made some contextual and semantic adaptations in relation to the others OSTRC-QHP versions due to the mean age (15) and the education level (70% incomplete secondary education) of our cohort. Therefore, we believe that the results from the clinimetric properties analyses reflect this population and may not be the same as those of other populations with different ages and education levels. Finally, although we expected the presence of floor effect, because we knew that during the follow-up there would be many times when athletes would not report health problems, this may be considered a limitation of this study.

Conclusion

This study provides an adequate and culturally adapted OSTRC-QHP questionnaire to be used in future research in countries that use the BrP language. The versatility of the OSTRC-BR, which can be adapted for various sports, and the simplicity of its application make the OSTRC-BR a very practical tool for athlete surveillance.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors would like to thank to Ronaldo Alves da Cunha, Marcelo Bannwart Santos, Beatriz da Silva Peres and all athletes from COTP for their support in this study. Luiz Hespanhol was granted with a Young Investigator Grant from FAPESP, grant 2016/09220-1.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjpt.2021.06.010.

References


